doi:10.1006/jmbi.2000.3512 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2000) 296, 833–849

JMB

Guided Selection of a Pan Carcinoma Specific Antibody Reveals Similar Binding Characteristics yet Structural Divergence Between the Original Murine Antibody and its Human Equivalent

Sigrid H. W. Belboer[1], Anneke Reurs[1], Rob C. Roovers[1] Jan-Willem Arends[1,2], Nick R. Whitelegg[3], Anthony R. Rees[3] and Hennie R. Hoogenboom[1]*

[1]*Research Institute Growth and Development, Department of Pathology, Maastricht University and*

[2]*University Hospital Maastricht, P.O. Box 616 6200, MD Maastricht, The Netherlands*

[3]*Department of Biology and Biochemistry, University of Bath, United Kingdom*

**Corresponding author*

Antibody engineering provides an excellent tool for the generation of human immunotherapeutics for the targeted treatment of solid tumours. We have engineered and selected a completely human antibody to epithelial glycoprotein-2 (EGP-2), a transmembrane glycoprotein present on virtually all human simple epithelia and abundantly expressed on a variety of human carcinomas. We chose to use the procedure of "guided selection" to rebuild a high-affinity murine antibody into a human antibody, using two consecutive rounds of variable domain shuffling and phage library selection. As a starting antibody, the murine antibody MOC-31 was used. After the first round of guided selection, where the V_H of MOC-31 was combined in Fab format with a human V_LC_L library, a small panel of human light chains was identified, originating from a segment of the VκIII family, whereas the MOC-31 V_L is more homologous to the VκII family. Nevertheless, one of the chimaeric Fabs, C3, displayed an off-rate similar to MOC-31 scFv. Combining the V_L of C3 with a human V_H library, while retaining the V_H CDR3 of MOC-31, clones were selected using human V_H genes originating from the rarely used V_H7 family. The best clone, 9E, shows over 13 amino acid mutations from the germline sequence, has an off-rate comparable to the original antibody and specifically binds to the "MOC-31"-epitope on EGP-2 in specificity and competition ELISA, FACS analysis and immunohistochemistry. In both V_L and V_H of antibody 9E, three germline mutations were found creating the MOC-31 homologue residue. Structural modelling of both murine and human antibodies reveals that one of the germline mutations, 53Y in V_H CDR2, is likely to be involved in antigen binding. We conclude that, although they may bind the same epitope and have similar binding affinity to the antigen as the original murine antibody, human antibodies derived by guided selection unlike CDR-grafted antibodies, may retain only some of the original key elements of the binding site chemistry. The selected human anti-EGP-2 antibody will be a suitable reagent for tumour targeting.

Keywords: human antibody; tumour targeting; phage selection; humanisation; epithelial glycoprotein-2

Present address, A. R. Rees, Syntem, Parc Scientifique Georges Besse, Nimes, France.

Abbreviations used: Ag, antigen; CDR, complementarity determining region; ELISA, enzyme-linked immunosorbent assay; Fab, antigen-binding fragment; FR, framework; mAb, monoclonal antibody; PBS, phosphate-buffered saline; scFv, single-chain Fv; SPR, surface plasmon resonance; V, variable; V_H, variable domain of the heavy chain; V_L, variable domain of the light chain; EGP-2, epithelial glycoprotein-2; TU, transforming units.

E-mail address of the corresponding author: hho@lpat.azm.nl

0022-2836/00/030833–17 $35.00/0

Introduction

By the expression of tumour-associated antigens (TAA), colorectal carcinoma (CRC) cells may be distinguished from their non-malignant counterparts. Among the identified surface-bound TAAs is epithelial glycoprotein-2 (EGP-2), also known as CO17-1A antigen, KSA, EGP40 or Ep-CAM. The antigen, encoded by the GA733-2 gene, is a 38 kDa transmembrane protein that is present on human non-squamous epithelia and is highly expressed on their derived tumours. In the serum of patients suffering from carcinoma, EGP-2 could not be detected by MOC-31 (de Jonge *et al.*, 1993), one of the many murine monoclonal antibodies (mAbs) that have been generated against the EGP-2 antigen. Thus, the EGP-2 antigen is not shed into the circulation and is regarded as a suitable target for imaging and immunotherapy of carcinomas. In radioimmune detection (Balaban *et al.*, 1991; Kosterink *et al.*, 1995), targeting of toxins or chemotherapeutic drugs (LeMaistre *et al.*, 1987; Zimmermann *et al.*, 1997; Elias *et al.*, 1994) and phase I and II clinical trials (Riethmuller *et al.*, 1994, 1998; Elias *et al.*, 1990), murine mAbs were successfully used. These studies show the large potential of mAbs to EGP-2 in immunotherapy.

In a previous study, we showed that the Mab MOC-31 has the lowest kinetic off-rate of a series of well-characterised anti-EGP-2 antibodies (Roovers *et al.*, 1998). The V-genes of this antibody have been cloned as single chain variable fragment (scFv) that competes with the original hybridoma antibody for binding to the antigen. The off-rate of the scFv is better than those of the bivalent 17-1A and 323/A3 anti-EGP-2 antibodies, providing it with an essential characteristic for tumour retention *in vivo* (Adams *et al.*, 1998). Therefore, the V-genes of MOC-31 are very suitable for the rational design and generation of human antibody-based immunotherapeutics for the treatment of solid tumours.

A disadvantage of the use of murine antibodies for immunotherapy of human carcinomas is that during the repeated administration of the mAbs to patients, a human anti-mouse antibody (HAMA) immune response can be induced. This may cause rapid blood clearance of the mAbs and serum sickness, thereby reducing their efficacy (Khazaeli *et al.*, 1994). It is expected that human antibodies would evoke a more effective and compatible response with the patient's immune system. Murine antibodies can be humanised by CDR-grafting (Jones *et al.*, 1986), where the hypervariable loops of the mAb are transplanted to a human antibody, thereby retaining the epitope and specificity. However, such CDR-grafted antibodies retain a large proportion of non-human sequences, which include the CDRs as well as framework residues that are sometimes retained because they are involved (in)directly in antigen binding (Foote & Winter, 1992; Baca *et al.*, 1997). Furthermore, humanisation by resurfacing (Roguska *et al.*, 1996) involves the retention of non-human framework sequences in all but surface located positions. In all instances, some rodent antibody-derived sequences are still included in the humanised antibody and thereby may enable the mAb to provoke an immune response. In particular, the transplantation of CDR-loops with "non-human" canonical forms, such as canonical structures 1 and 5 of murine light chain loop L1 (Tomlinson *et al.*, 1995) for which there is no human homologue, may assign a typical murine idiotype structure to the antibody-combining surface, as well as provide unique T-cell epitopes for HAMA induction.

An alternative method, termed guided selection, has been developed to convert murine antibodies into completely human antibodies with similar binding characteristics (Jespers *et al.*, 1994). In this method, the murine V domains are sequentially or in parallel replaced by human V domains, using phage selection to derive the human antibody with the best affinity. Two published examples demonstrate the power of this method. Figini *et al.* (1994) shuffled the light chain of a murine anti-hapten phenyl-oxazolone antibody, NQ10.12.5, with a repertoire of human heavy chains. After a second shuffle of the selected V_Hs with a repertoire of human V_Ls, entirely human antibodies were generated. In a similar approach, starting with the murine V_H, a murine anti-TNFα (tumour necrosis factor) antibody was converted into a human version (Jespers *et al.*, 1994).

The combination of shuffling of V-genes and selection on antigen provides a means to direct the isolation of a human V-gene pair with largely similar binding characteristics as the starting antibody. In order to obtain a human anti-EGP-2 antibody with the same binding characteristics as MOC-31, which may be used for immunotherapy, we chose the method described by Jespers *et al.* (1994) for the humanisation of the MOC-31 mAb by guided selection. In two sequential chain shuffles, the murine V_L and V_H were replaced by fully human antibody V genes, while maintaining both specificity and affinity. The murine and fully human antibodies were modelled and structural analogies and differences are discussed.

Results

Humanisation of the light chain

In the first guided selection step, we aimed to replace the murine light chain with a human version. The variable heavy chain gene of MOC-31 was cloned into the pCES1 Fab format vector (de Haard *et al.*, 1999) and the obtained chimaeric V_HC_H1 was combined with κ and λ light chain libraries derived from human spleen separately, yielding two libraries of 10^8 clones (κ) and 2×10^7 clones (λ). We chose to use spleen as a source of human V-genes to access maximal germline and somatic hypermutation diversity (of both IgM and IgG B cells). Phage were rescued and 8×10^{12}

transforming units (TU) of the κ library were used for the first selection round. Schier *et al.* (Schier & Marks, 1996; Schier *et al.*, 1996) obtained higher-affinity monomeric scFv fragments on phage only by selecting in solution using limiting concentrations of biotinylated antigen, followed by screening the selected binders for their k_{off} in BIAcore. In order to obtain high-affinity antibodies to tumour antigen EGP-2, selections were performed on biotinylated EGP-2 with concentrations varying from 1 nM down to 10 pM (Table 1). After each round of selection, polyclonal phage was prepared and the frequency of binders to EGP-2 was determined in ELISA. The ELISA results in Table 1 show that in all selection rounds binders were retrieved. Clones positive in ELISA and selected in the third round, were tested for their affinity in BIAcore: some Fabs show an off-rate similar to the starting antibody MOC-31 (Table 2). The V_L-gene nucleic acid sequence of these clones was determined and submitted to GenBank (accession numbers: AF209862-AF209866). Alignment using V BASE, revealed that the light chains originate from the human VκIII family (Figure 1). No correlation can be found between the antigen concentration used for selection and the off-rate of the selected clones. For instance, clones C1, C5 and E4, which by sequencing appeared identical, were selected using an EGP-2 concentration of 1 nM (C1) and 0.1 nM (C5) or a combination of these (E4, see Table 2). To produce more reliable off-rates, the Fabs with the lowest off-rates were produced in large amounts, purified and analysed in BIAcore. As shown in Table 2, the purified Fabs show an off-rate in the same range as MOC-31 (between 10^{-4} s^{-1} and 10^{-3} s^{-1}). Clone C3 was chosen to guide the selection of the human heavy chain. Compared to the sequence of the germline segment, C3 shows 16 nucleotide and nine amino acid substitutions.

As shown in Table 3, the selected human V_Ls have a different canonical structure for L1 (fold 2) to that used by MOC-31 V_L (fold 4). No differences in lengths or canonical loops were observed for CDR2. Three of the selected light chains used the same canonical structure for L3 (fold 1) as MOC-31. The structure for L3 of the other selected V_Ls (having one extra residue in L3) could not be determined, since these chains have threonine at position 96, whereas fold 5 (with seven residues in the L3 region) uses proline at this position (Tomlinson *et al.*, 1995). Furthermore, no significant amino acid homology was found between the CDRs of the human V_Ls and those of the MOC-31 V_L (Figure 1).

In parallel with the selections with the κ library, selections were performed with the λ library. No enrichment could be found after three rounds of selection and no EGP-2 binding clones were detected in ELISA. Most likely, for replacement of the murine κ chain, there was no V_H or antigen binding-compatible λ light chain available in the human λ repertoire.

Humanisation of the heavy chain

For the chain shuffling of the heavy chain, the selected human light chain of clone C3 was first cloned into pCES1 not carrying any V_H genes to avoid later contamination of the MOC-31 V_H. Human V_H genes were amplified from the RNA of human spleen using oligonucleotide H3-MOC-31-FOR, introducing the MOC-31 CDR3 in all sequences. This product was cloned as indicated in Materials and Methods into the C3-light chain containing vector, to create a library of 1.2×10^7 clones. This was subjected to three rounds of selection. Under stringent selection conditions (decreasing phage input titres and decreasing antigen concentration, as before), no human EGP-2 binding antibodies were identified. After three rounds of selection, EGP-2 binding antibodies were obtained using non-stringent conditions (phage input 10^{12} TU, 100 nM antigen). Figure 2 shows that all 11 clones tested are specific for EGP-2, but signals do not reach values above 0.7. DNA sequencing of four EGP-2 specific clones (including clone 9E) revealed that these clones all contain the same V_H. This V_H was completely human and its DNA sequence was submitted to GenBank (accession number: AF209860). V BASE sequence alignment showed it originated from the small but discrete V_H7 family (Willems van Dijk *et al.*, 1993) with 23 nucleotide and 13 amino acid mutations compared to the germline sequence (Figure 3). As shown in Figure 3, the amino acid sequence in the CDR3 and framework 4 regions of MOC-31 VH were retained during the shuffling.

BIAcore experiments with the representative clone, 9E, purified from 0.5 litre of culture, show that it has an affinity of the same order of magnitude as MOC-31 (Table 2). The human Fab 9E

Table 1. Binding frequency of selected Fabs from κ library

[Ag] used during selection	Round of selection			
(nM)	1	2	3	4
1 (rd1); 1 (rd2); 1 (rd3); 1 (rd4)	4/30	0/8	12/16	nd
1 (rd1); 0.1 (rd2); 0.1 (rd3); 0.1 (rd4)	4/30	0/8	11/16	nd
0.1 (rd1); 0.1 (rd2); 0.1 (rd3); 0.1 (rd4)	0/32	0/8	7/16	12/16
0.1 (rd1); 0.01 (rd2); 0.01 (rd3); 0.01 (rd4)	0/32	0/8	6/16	0/8

nd, not determined; rd, round.

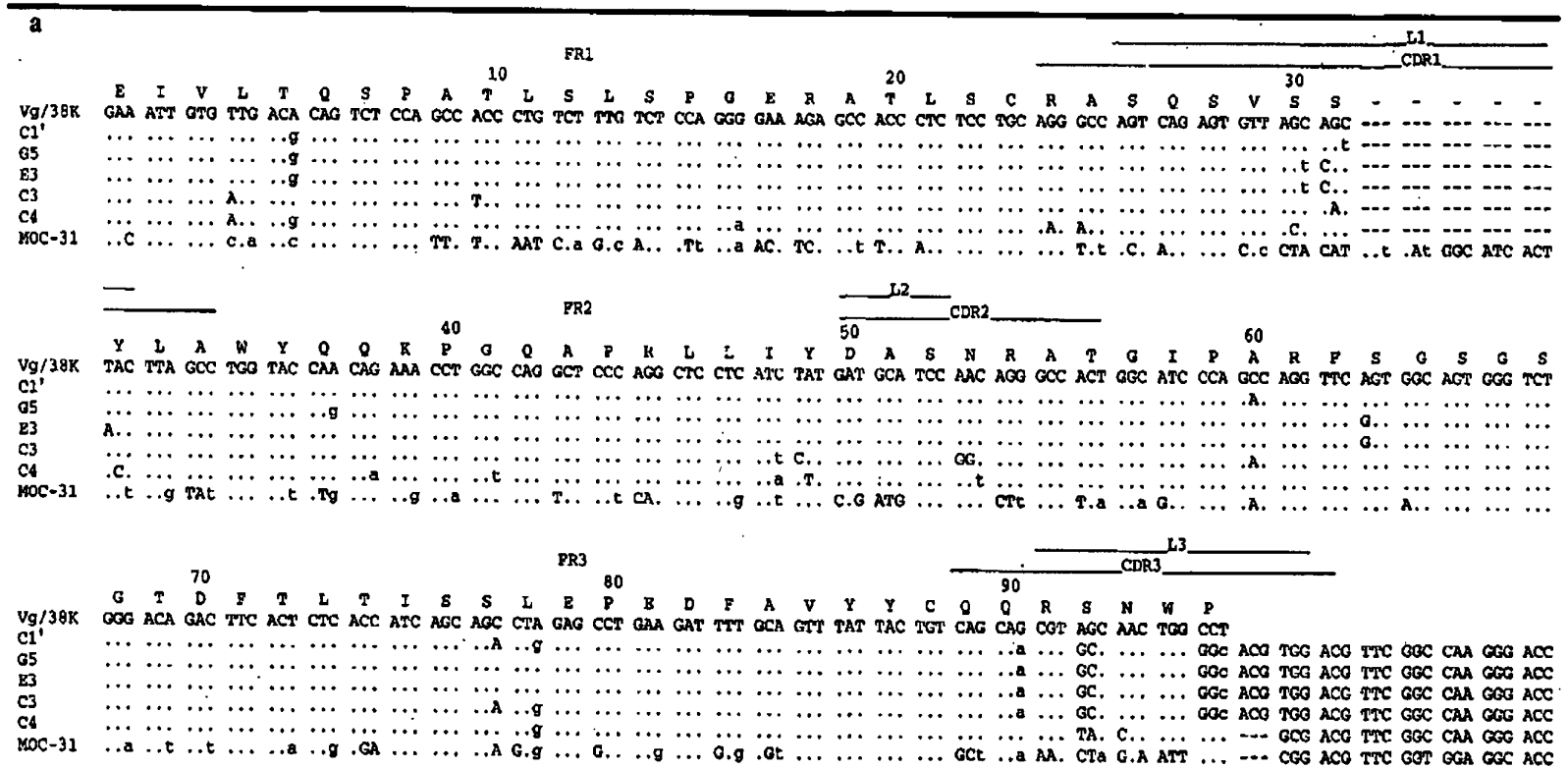


Figure 1 (*legend opposite*)

b

	FR1			CDR1			FR2			CDR2	
	10		20	30			40			50	
Vg/38K	EIVLTQSPATLSLSPGERATLSC			RASQSVSS-----YLA			WYQQKPGQAPRLLIY			DASNRAT	
C1'	.....			.....------...			.....			.....	
G5	.....			.....R-----...			.....			.....	
E3	.....			.....R-----N..			.....			.....	
C3	...M.....S.....			.....N-----...			.....H			...G...	
C4	...M.....			KT....T.-----S..			.....P			.....	
MOC-31	D.....FSNPVTL.TS.SI..			.STK.LLHSNGIT..Y			..L.....S.Q....			QM..L.S	

	FR3			CDR3	FR4
	60	70	80	90	
Vg/38K	GIPARFSGSGSGTDPTLTISSLEPEDFAVYYC			QQRSNWP	
C1'	...D.....R.....			...A..GTWT	FGQGTKVEIKR
G5	.....G.....			...A..GTWT	FGQGTKVEIKR
E3	.....G.....			...A..GTWT	FGQGTKVEIKR
C3	...D.....R.....			...A..GTWT	FGQGTKVEIKR
C4	.....			...YH..-AT	FGQGTKVEIKR
MOC-31	.V.D...S.....R..RV.A..VG....			A.NLEI.-RT	PGGGTKLEIKR

Figure 1. DNA and amino acid sequence alignment of selected human variable light chains and MOC-31 V_L. (a) DNA sequence; (b) amino acid sequence. The DNA sequences of the V_L of half human Fabs (mouse V_H, human V_L) and of MOC-31 are aligned to Vg/38 K, the human germline gene with closest DNA sequence to the selected human light chains as deduced using V BASE (http://www.mrc-cpe.cam.ac.uk/imt-doc/restricted/DNAPLOT.html). Framework (FR) and CDR regions are indicated. Numbering and loop regions (L1-L3) are according to structural criteria defined by Chothia (Chothia *et al.*, 1992; Tomlinson *et al.*, 1992, 1995; Williams *et al.*, 1996). CDRs are according to Kabat *et al.* (1991). Silent mutations from germline are indicated using small characters. † The DNA and amino acid sequence of clone C1 are identical with those of clones C5 and E4.

specifically binds EGP-2 positive cells as shown by FACS analysis (Figure 4) and competes with MOC-31 for binding to recombinant EGP-2 in ELISA (Figure 5). The Fab was also tested in immunohistochemistry and a similar staining was observed as shown by Roovers *et al.* (1998): the Fab reacts with normal and malignant non-squamous epithelia and does not stain malignant tissue reported negative for the antigen (melanoma; data not shown; Herlyn *et al.*, 1979). This confirms the specificity. The variable chains of MOC-31 and 9E were modelled as described in Materials and Methods. The coordinates of the V-genes of MOC-31 and 9E were deposited in the PDB-database with the accession numbers 1dx2 and 1dx3, respectively. As shown in Figures 6 and 7 and discussed in more detail in Discussion, the structural modelling revealed large differences between the murine and the human antibodies, in particular between their light chains.

Discussion

We cloned the repertoire of human light chains into the phagemid pCES1 containing the MOC-31

Table 2. On and off-rates of (half) human Fabs and scFv MOC-31

Clone[a]	[Ag][b] (nM) 1st rd	2nd rd	3rd rd	k_{off} (10^{-4} s^{-1})	k_{off} (10^{-4} s^{-1})[c]	k_{on}(10^5 M^{-1} s^{-1})[c]
C1	1	1	1	1.5	nd	
C5	0.1	0.1	0.1	1.6	nd	
E4	1	0.1	0.1	1.2	4.4	
G5	0.1	0.1	0.1	1.8	3.6	
E3	1	0.1	0.1	1.8	2.1	
C3	1	0.1	0.1	0.84	3.0	0.90
C4	1	0.1	0.1	34	nd	
9E[d]	100	100	100	nd	4.4	0.86
MOC-31[e]	na	na	na	nd	5.7	1.2

[a] Clones selected in 3rd round.
[b] The Ag concentration used in the selection round 1, 2 and 3.
[c] Off-rate of purified Fabs and scFv as determined by BIAcore.
[d] Fab 9E is completely human.
[e] scFv MOC-31 is completely murine.
na, not applicable; nd, not determined; rd, round.

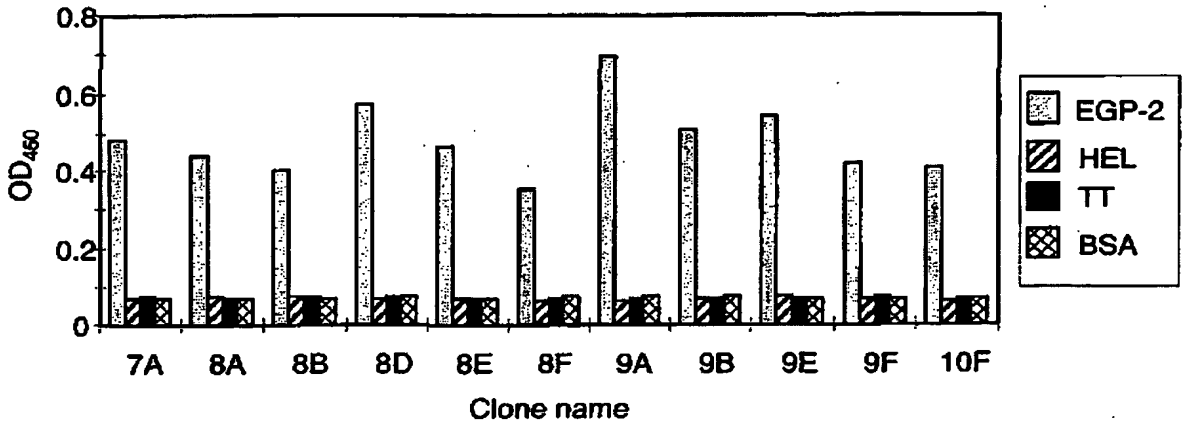


Figure 2. Specificity ELISA of clones selected after second chain shuffle. The specificity of the selected Fabs was tested in an ELISA using plastic-immobilised bovine serum albumin, hen egg-white lysozyme, tetanus toxoid and EGP-2. Bound soluble Fab fragments from individual clones were detected with anti-myc tag antibody 9E10.

V_H. Since MOC-31 scFvs have a tendency to dimerise (Roovers *et al.*, 1998), we used the Fab format for the humanisation. In this way, we can screen the antibodies for affinity rather than avidity and also select with less avidity-mediated problems for high affinity binders. We selected on soluble, biotinylated antigen using stringent selection conditions (low Ag concentration and low phage input). These selections yielded chimaeric murine V_H/human V_L antibodies with comparable off-rates as the parent scFv MOC-31. In contrast to the rapid and successful chain shuffling of the light chain, the shuffling of the heavy chain was more difficult: only by increasing the antigen concentration and the phage input could binders be selected. The lower expression yields of the fully human Fab 9E (34 μg/l) *versus* the Fab C3 (80 μg/l) as well as its reduced affinity are most likely responsible for the higher antigen concentration required to select the phage antibody. This emphasises the importance of the empirical approach for determining optimal conditions for affinity-selections.

Sequence alignment using V BASE showed that MOC-31 V_L has more homology to the VκII family than to the VκIII family from which the selected human light chains originated. Thus, unlike other examples (Figini *et al.*, 1994; Salfeld *et al.*, 1997), the light chain guided selection did not yield the most homologous human counterpart sequence. Nevertheless, there are signs of amino acid conservation between the MOC31 V_L and the selected human V_Ls. Remarkably, the most abundant clones, C1, C5 and E4, which are selected on different antigen

Table 3. V-gene classifications and structural predictions of the CDR-loops of V-genes

V-gene	Human V-gene family[a]	No. of mutations from germline[b]		Canonical structure[c]			Closest human[a] germline gene
		Nucleotides	Amino acids	L1	L2	L3	
C1[d]	VκIII	11	4	2	1	?	3A9
C5	VκIII	11	4	2	1	?	Vg/38K
E3	VκIII	11	5	2	1	?	Vg/38K
C3	VκIII	16	9	2	1	?	3A9
C4	VκIII	16	8	2	1	1	Vg/38K
VL MOC-31	VκII	109	44	4	1	1	DPK13/O11
9E	V_H7	23	13	1	2	NA	VI-4.1b
VH MOC-31	V_H7	70	26	1	2	NA	YAC-10/1d37

[a] V-gene families and closest human germlines (for MOC-31) were deduced from V BASE.

[b] Germline genes used are Vg/38K for the light chains and VI-4.1b for heavy chains.

[c] Canonical structures were assigned according to Chothia *et al.* (1989, 1992). Selected human V_Ls were assigned in combination with MOC-31 V_H. Selected human V_H (9E) was assigned in combination with the V_L of C3.

[d] The DNA sequence and therefore the V-gene classifications and structural predictions of CDR-loops of clone C1 are identical with clones C5 and E4.

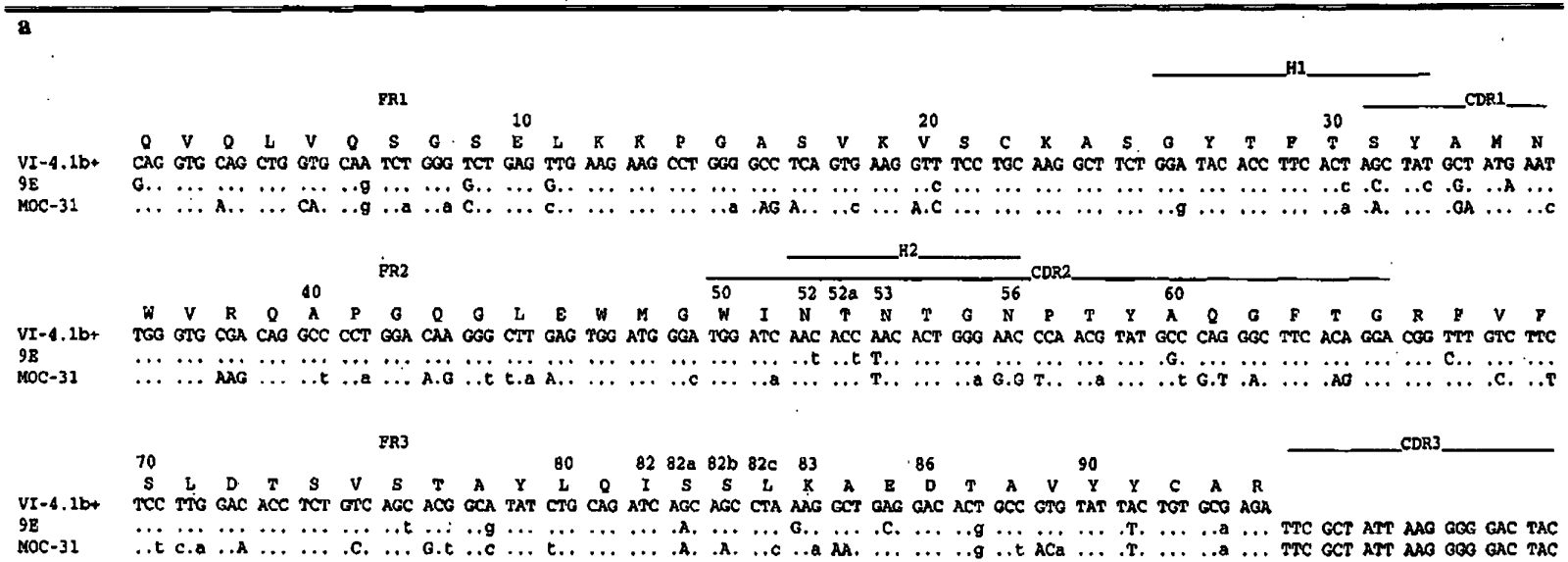
? No canonical of the same loop length could be found using webpage http://www.biochem.ucl.ac.uk/~martin/abs/chothia.html. But according to Tomlinson *et al.* (1995) canonical structure 5 might be assigned to L3 of these clones.

concentrations and are identical at the nucleotide level, are completely germline in sequence except for two mutations in FR3. At both sites, the germline residues are replaced by equivalent residues present in the original MOC-31 V_L sequence (A60D; S77R; Figure 1). The V_L of another selected clone, C3, shows the same mutations as well as a similar type mutation in FR1 (T10S). Of all selected human light chains, the V_L of clone C3 has both the highest level of amino acid substitutions from germline (Table 3) and the highest level of amino acid homology to MOC-31 V_L. This may well indicate the involvement of some of these residues in antigen binding. Structural modelling of scFvs MOC-31 and 9E (which carries the C3 light chain; see Figure 6) revealed that 60D is quite close (6 Å) to 54R in the L2 loop, but faces the other way and hence cannot form a salt bridge with it. The conformation of its side-chain, however, is similar to that in MOC-31. The other germline mutations are in the same conformation in both MOC-31 and 9E and appear to be largely unimportant, being framework residues distant from the CDRs which, additionally, point sideways or interact only with the framework. As shown in Figure 6(c) and (d), the exposed residues 24R, 27K, 31H, 93E present in the MOC-31 V_L were not retained in C3-V_L except for 24R, but this residue is projected from the side of the Fv and is therefore unlikely to be involved in antigen binding. In the humanised C3 light chain the exposed residues are 24R, 50D, 54R, 91R and to some extent 28S, 30S and 32Y. Residue 54R is projected downwards and sideways, away from the binding site (not shown in Figure 6). When we compare the side views of both Fvs, we observe the bulky residue 94W in the centre of 9E (see Figure 6(b) and (d)), which is absent in MOC-31 (Figure 6(a) and (c)). These observations suggest that for the light chain guided selection, sequence homology is not required to maintain an antigen-antibody interaction with the same or similar kinetics of binding. This may be caused by the often-dominant role of the heavy chain in antigen binding.

Sequence alignment using V BASE of the V_H of MOC-31 showed that this chain has the highest homology with germline YAC-10/1d37 of the human V_H7 family. Indeed, the selected human heavy chain of clone 9E originated from the V_H7 family, albeit from a different germline segment (VI-4.1b). It differs only moderately from the murine MOC-31 V_H (Table 3), in total in 29% (26/90 residues, in which the 90 are encoding the region after the FR1-part encoded by the PCR-primers and up to and including FR4). The antibody models were used to investigate possible common features around the antibody combining site. The main structure of the heavy chain CDR3 was, as expected, maintained. For example, from Figure 6(e) and (f) it is clear that the heavy chain residues 98K and 100D are buried in both antibodies. Of the 13 amino acid germline mutations found in 9E, three are mutations that create the MOC-31 homologue residue (A33G in CDR1, N53Y in CDR2 and S82aN in FR3). As shown in Figure 6, of the exposed residues 53Y, 56E and to some extent 28T in MOC-31 V_H, only 53Y is retained in 9E-V_H, suggesting the involvement of this residue in antigen binding. The superimposition of the binding sites of both antibodies, shown in Figure 7, reveals that some residues have very similar positions and that they are actually in corresponding positions in the CDR loops. In particular, this is very obvious for the V_H residues at positions 53 and 56. The pattern of mutations and the fact that a non-homologous human V_L and a human heavy chain with close homology to its MOC-31 counterpart were selected, suggest that the heavy chain of MOC-31 is more essential for the specific binding of the antibody to EGP-2 than the light chain. Thus, the guided selection procedure yields a human antibody with similar binding characteristics as the original murine MOC-31 antibody, yet with minimal structural conservation. This contrasts with the data from Figini *et al.* (1994), in which an anti-phenyl-oxazolone antibody was humanised with the same procedure; in this case, six/seven of the putative contact residues were retained in the fully human version. The structural solution to the recognition of a particular epitope within a larger protein antigen is less stringent. For example, minimal structural homology has been noted in two anti-influenza virus neuraminidase antibodies binding to overlapping epitopes (Malby *et al.*, 1994). This reduced stringency on conservation of critical binding site residues within the antibody-protein antigen interaction, apparently allows structural divergence of antibodies obtained by guided selection.

Using the humanisation technique of guided selection, we have isolated a fully human antibody directed to EGP-2. Since 1994, several groups have used the guided selection technique to obtain human antibodies by sequential chain shuffling (Salfeld *et al.*, 1997; Figini *et al.*, 1998; Rader *et al.*, 1998; Watzka *et al.*, 1998). Watzka *et al.* (1998) reported that the second shuffle (of the V_H) led to the selection of a humanised antibody, which recognised an epitope present only on the immobilised (coated) antigen. In principle, the modulation of fine-specificity may be caused by the possible different chemical interactions with a similar area on the antigen. However, in this case it appears that a suboptimal selection procedure has interfered with the structure of the epitope, causing failure of the guided selection procedure, and selection of antibodies to completely different epitopes on the antigen. Since the heavy chain CDR3 is the main loop involved in antigen binding and its species origin cannot be traced easily due to its highly diverse nature, several groups have opted to maintain this loop in the guided selection pro-

a



b

	FR1 (10–30)	CDR1	FR2 (40)	CDR2 (50 52a 56 60)
VI4.1b	QVQLVQSGSELKKPGASVKVSCKASGYTFT	SYAMN	WVRQAPGQGLEWMG	WINTNTGNPTYAQGFTG
9E	E.....A.V.....	T.GI.	.....	....Y.....G.....
MOC-31	..K.Q...P.....ET..I.....	N.G..	..K....K..K...	....Y..ES...DD.K.

	FR3 (70–90)	CDR3	FR4
VI4.1b	RFVFSLDTSVSTAYLQISSLKAEDTAVYYCAR		
9E	.L.....N..E.A.....F...	FAIKGDY	WGQGTLVTVSS
MOC-31	..A...E..A.A.....NN..N....T.F...	FAIKGDY	WGQGTTVTVSS

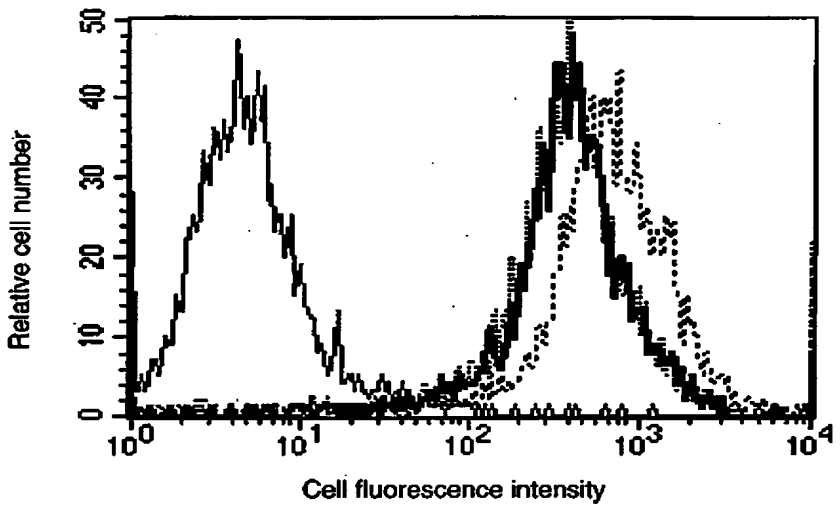


Figure 4. Binding of soluble scFv MOC-31 and Fabs C3 and 9E to $CaCO_2$ cells in flow cytometry. Thin line is negative control (9E10); thick line is binding of scFv MOC-31; (...) half human Fab C3; (---) human Fab 9E.

cedure. Similarly, we have succeeded in humanising an anti-CD30 antibody, which had the same epitope specificity as the original murine Ki-4 antibody, by retaining only the murine heavy chain CDR3 (Klimka *et al.*, unpublished results). Rader *et al.* (1998) humanised the murine antibody LM609 (directed to human integrin $\alpha_v\beta_3$) by guided selection, retaining the murine CDR3s of both heavy and light chain. In contrast, our data suggest that it may not always be necessary to maintain the CDR3 of the murine light chain during the guided selection, as was suggested by Rader *et al.* (1998). Thus, guided selection of rodent antibody V-genes provides a powerful tool for interspecies conversion of antibodies. Due to the different chemistry of the human antibody-antigen interaction, it can be envisaged that in some cases guided selection may alter the antibody-mediated triggering or signal transduction of the starting rodent antibody (i.e. antibody internalisation, receptor cross-linking). Therefore, any resulting antibody will need to be screened for loss (or gain) of such secondary binding feature.

In conclusion, we have obtained a high-affinity antibody that specifically binds EGP-2 as confirmed by FACS analysis, immunohistochemistry and specificity ELISA and may be useful for tumour targeting studies. For instance, bispecific antibodies can be generated by combining the isolated human α-EGP-2 antibody with an α-CD3 antibody. Together with co-stimulatory signals these bispecific antibodies can be used to provide T cells with tumour specificity and cytotoxic potential (Kroesen *et al.*, 1994). Recently, such a bispecific molecule was synthesised: the V-genes of MOC-31 were combined with those of an α-CD3 antibody. This diabody was capable of retargeting T cells to lung cancer cells *in vitro* (Helfrich *et al.*, 1998). Thus, the human α-EGP-2 antibody may be suitable as a building block for the generation of immunotargeting therapeutics such as immunotoxins, immunocytokines or bispecific, whole or radiolabelled antibodies.

Materials and Methods

Escherichia coli strain

Escherichia coli TG1: K12, D(*lac-pro*), *supE*, *thi*, *hsdD5/F' traD36*, *proA*$^+$*B*$^+$, *lacI*^{q}, *lacZ*DM15.

Figure 3. DNA and amino acid sequence alignment of the selected human variable heavy chain and MOC-31 V_H. (a) DNA sequence; (b) amino acid sequence. The DNA sequence of the human heavy chain from the selected anti-EGP-2 clone and of MOC-31 are aligned to VI-4.1b, the human germline gene with closest DNA sequence to the selected human heavy chain as deduced using V BASE (http://www.mrc-cpe.cam.ac.uk/imt-doc/restricted/DNA-PLOT.html). Framework (FR) and CDR regions are indicated. Numbering and loop regions (H1-H3) are according to structural criteria defined by Chothia (Chothia *et al.*, 1992; Tomlinson *et al.*, 1992, 1995; Williams *et al.*, 1996). CDRs are according to Kabat *et al.* (1991). Silent mutations from germline are indicated using small characters.

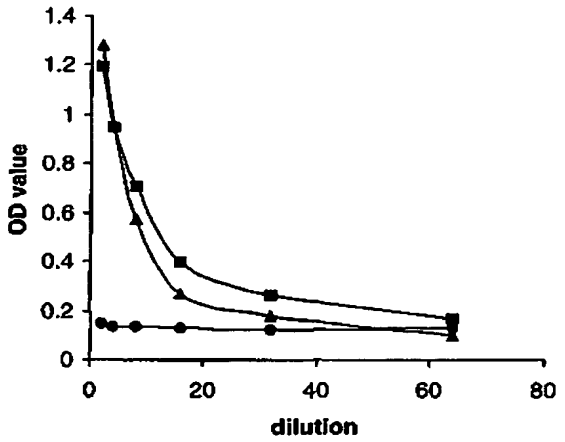


Figure 5. Competition ELISA. Binding of Fab 9E phage antibody to recombinant EGP-2 in the presence of (▲) no competing antibody, (■) an excess of an irrelevant antibody (UIC2, anti P-glycoprotein hybridoma) or (●) the MOC-31 whole antibody.

Humanisation of light and heavy chains

The V-genes of MOC-31 were cloned in the phagemid pCES1 Fab format vector (de Haard *et al.*, 1999). The *ApaLI-AscI* fragment of the cloned V_LC_L was replaced by the V_LC_Ls from a library (9×10^6 clones), which was prepared from total RNA isolated from a human spleen using random primers (Promega) for the preparation of cDNA. For the 2nd round of chain shuffling, the selected human V_LC_L was combined with a human V_H library. To retain the heavy chain CDR3 of MOC-31 the H3 MOC-31-for-primer (5'-ACT CGA GAC GGT GAC CAG GGT ACC TTG GCC CCA GTA GTC CCC CTT AAT AGC GAA TCT TGC ACA G(AT)A ATA CAC GGC CGT GTC-3') was used with an equimolar mixture of VHBACKSfi primers (de Haard *et al.*, 1999) for PCR amplification of the V_H genes from cDNA derived from a human spleen (de Haard *et al.*, 1999). The cDNA was prepared from total RNA using random primers (Promega) as described (de Haard *et al.*, 1999). The PCR products were cloned as a *SfiI-BstEII* fragment in pCES1 containing the selected human light chain and a human gamma-1 C_H1 gene fused to filamentous phage gene III, as described by de Haard *et al.* (1999).

Affinity selection

Recombinant EGP-2 was expressed in the baculovirus system as described (Strassburg *et al.*, 1992; a kind gift from Professor D. Herlyn (the Wistar Institute)) and was biotinylated with NHS-SS-biotin (Pierce) according to the manufacturer's descriptions. The biotinylated EGP-2 was tested with MOC-31 on a streptavidin-coated sensorchip in BIAcore2000 (Biacore AB, Sweden). Selections were performed on biotinylated EGP-2 as described by Hawkins *et al.* (1992) with some modifications: phage were incubated on a rotator wheel for one hour in 2% M-PBST (PBS, supplied with 2% (w/v) skimmed milk powder and 0.1% (v/v) Tween-20). Meanwhile, 100 µl of Streptavidin-conjugated paramagnetic beads (Dynal, Oslo, Norway) were incubated on a rotator wheel for two hours in 2% M-PBST. Different amounts of biotinylated EGP-2 were added to the pre-incubated phage (for the V_L-shuffle: 10^{12} TU for round 1, 10^{11} TU for subsequent rounds; for the V_H-shuffle: 10^{12} TU) and incubated on a rotator wheel for 30 minutes. Next, beads were added and the mixture was left on the rotator wheel for 15 minutes. After 14 washes with 2% M-PBST and one wash with PBS, phage particles were eluted with 950 µl of 0.1 M triethylamine for five minutes. The eluate was neutralised by the addition of 0.5 ml of Tris-HCl (pH 7.5) and was used for infection of log-phase *E. coli* TG1 cells. The TG1 cells were infected for 30 minutes at 37°C and were plated on 2xTY (16 g Bacto-trypton, 10 g Yeast-extract and 5 g NaCl per litre) agar plates, containing 2% (w/v) glucose and 100 µg/ml ampicillin. After overnight incubation at 30°C, the colonies were scraped from the plates and used for phage rescue as described (Marks *et al.*, 1991).

ELISA and kinetic measurement using SPR in BIAcore

Soluble scFvs and Fabs were produced as described (Roovers *et al.*, 1998).

ELISAs and kinetic measurements using SPR in a BIAcore™2000 instrument (BIAcore AB, Uppsala, Sweden) were performed as described by Roovers *et al.* (1998). For the specificity ELISA, wells were coated with 10 µg/ml bovine serum albumin (Sigma) in PBS, 3 mg/ml hen egg-white lysozyme (Boehringer Mannheim) in 0.1 M $NaHCO_3$ (pH 9.6), 10 µg/ml tetanus toxoid in 0.1 M $NaHCO_3$ (pH 9.6), or with 0.5 µg/ml EGP-2 in PBS. For the competition ELISA, Fab antibodies expressed as pIII fusions on the tip of bacteriophage were detected in the presence of excess whole antibody as described (Roovers *et al.*, 1998). Approximately 10^{13} colony-forming units (cfu) of phage (and twofold serial dilutions thereof) were mixed with 50 µl of hybridoma supernatant containing approximately 1 µg/ml whole antibody. For affinity measurements, the non-biotinylated antigen was covalently coupled to a CM-5 sensorchip (BIAcore AB, Uppsala, Sweden), resulting in a surface of 500 resonance units (RU). All kinetic measurements were performed on this antigen surface. Using TBS (20 mM Tris (pH 7.4), 150 mM NaCl) containing 2 mM $MgCl_2$ and 0.01% (v/v) Tween-20 as a running buffer, the Fabs/scFvs were passed over the sensorchip at a flow rate of 20 µl/minute at 25°C. The rate constant (k_{off}) was obtained by direct fitting according to a single-site model, using the BIAevaluation 2.1 software (BIAcore AB, Uppsala, Sweden). The off-rates (k_{off}) of 9E and MOC-31 are, respectively, an average of eight and five curves made with different antibody dilutions. Association rates (k_{on}) were determined from a secondary plot of ($k_s = k_{on} \times C + k_{off}$) *versus* antibody concentration (C).

DNA sequence analysis

The nucleotide sequences of the selected Fabs were determined using the dideoxy sequencing method of Sanger. Products of the sequencing reaction were ana-

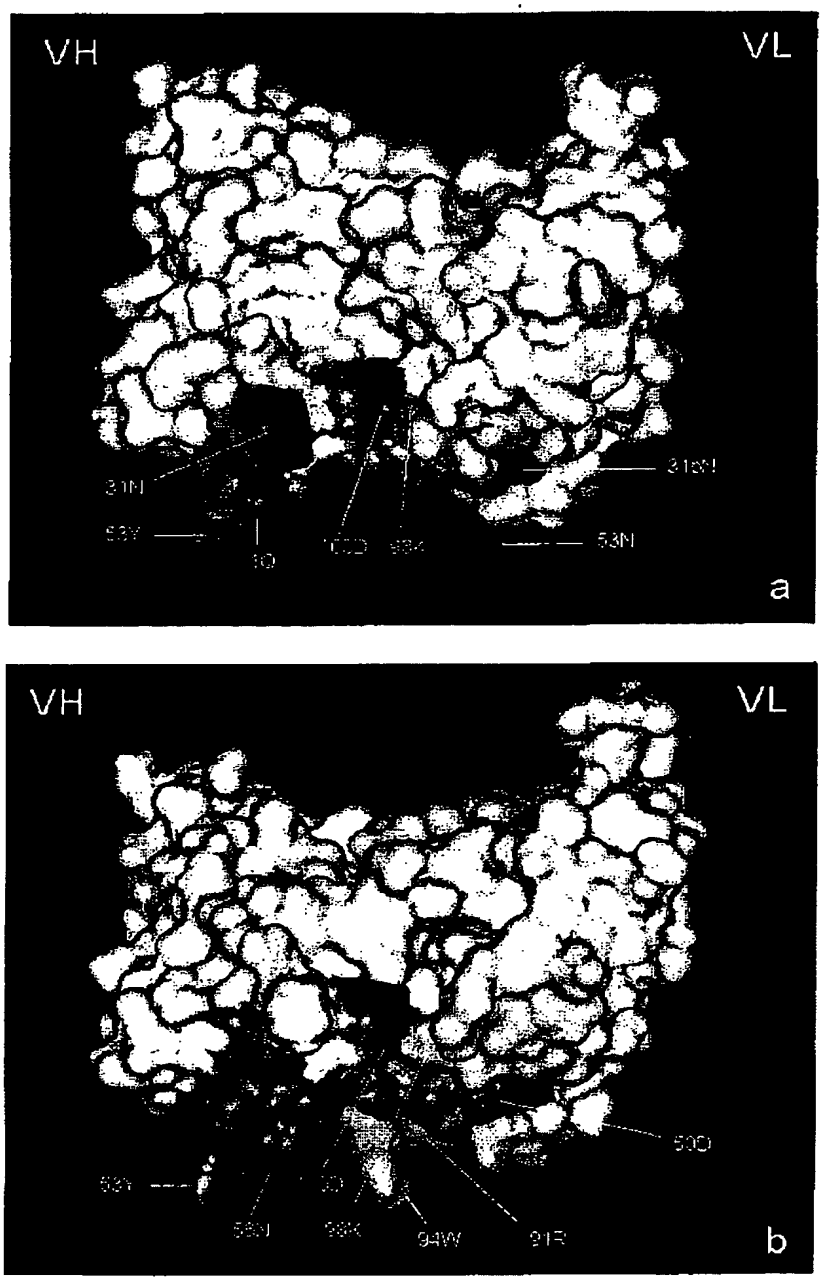


Figure 6 (*legend shown on page 833*)

lysed on a semi-automated sequencer (Alf Express, Pharmacia). Oligonucleotides used were pUC-REVERSE: 5′-CAG GAA ACA GCT ATG AC-3′ and CH1FOR: 5′-GTC CTT GAC CAG GCA GCC CAG GGC-3′. Alignments of V gene sequences were performed using V BASE on webpage http://www.mrc-cpe.cam.ac.uk/imt-doc/restricted/DNAPLOT.html. Canonical class assignments were performed using

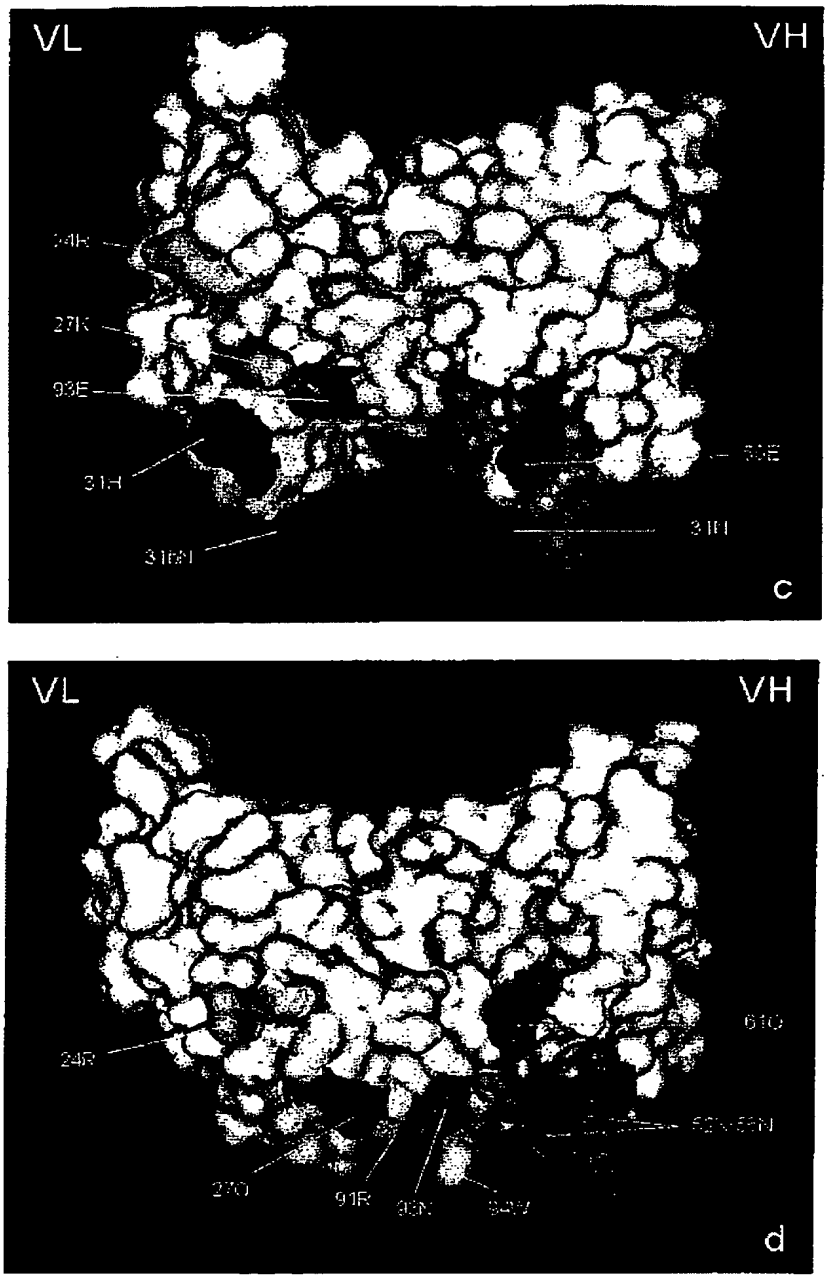


Figure 6 (*legend shown on page 833*)

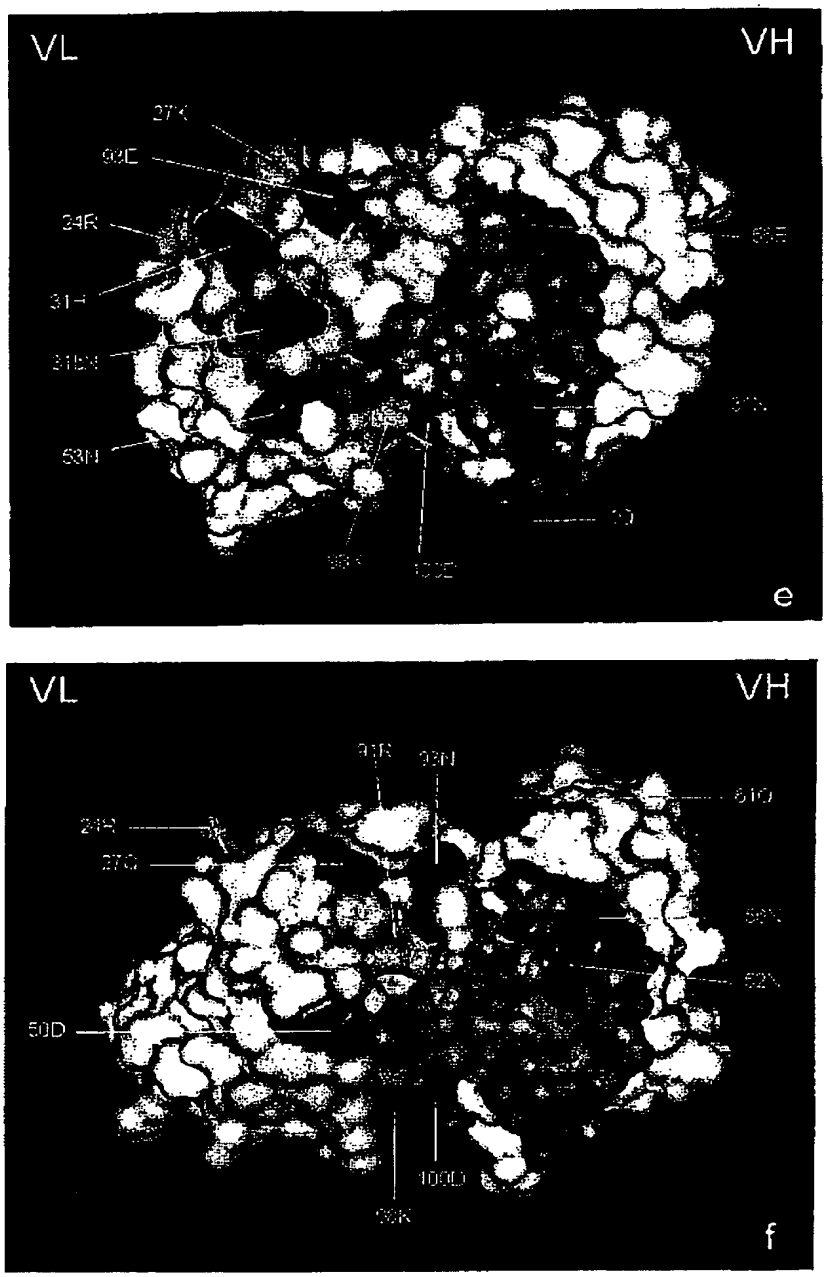


Figure 6. Connolly solid-surface structures of the monoclonal MOC-31 and the human Fab 9E. The variable chains only are shown. The CDRs of V_H are coloured orange and those of V_L are yellow. Positively charged residues (Arg and Lys) are coloured cyan and negatively charged residues (Asp and Glu) are coloured red. Asn and Gln residues present in CDRs are coloured magenta. The MOC-31 light chain histidine at position 31 is coloured blue. (a) and (b) The opposite side view of (c) and (d), respectively. (e) and (f) Top views showing the antigen binding site. (a), (c) and (e) MOC-31; (b), (d), and (f) 9E. The Connolly structures were generated using the program Insight-II 97.0 (MSI) and show the solvent-accessible, rather than the van der Waal surface.

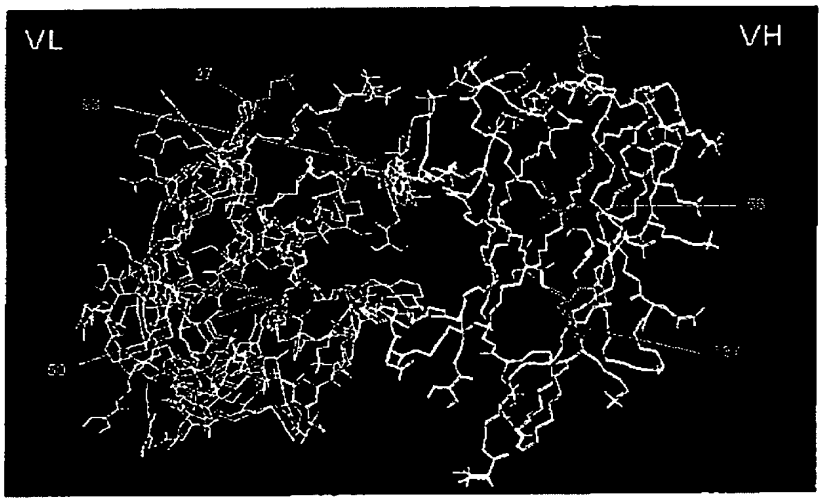


Figure 7. Superimposition of the antigen binding sites of the monoclonal MOC-31 and the human Fab 9E. The variable chains only are shown. The two models were superimposed using least-squares fitting on conserved framework regions adjacent to the CDRs. Residues with similar positions are shown in red (9E) or green (MOC-31).

webpage http://www.biochem.ucl.ac.uk/~martin/abs/chothia.html.

Structural modelling

The antibody modelling software that incorporates the following methods is described in detail on the webpage http://antibody.bath.ac.uk

Building the framework and canonical CDRs

The framework and canonical CDR backbones were modelled by an improved version (N.R.W. & A.R.R., unpublished) of the original AbM algorithm (Martin *et al.*, 1989). As a result of studies which showed that certain residue positions in canonical loops were conserved in chi-1 angles, the side-chains of those positions were built using maximum-overlap, rather than using the CONGEN iterative side-chain placement algorithm (Bruccoleri & Karplus, 1987). The non-conserved side-chains were not built at this point, but were built simultaneously with the non-conserved H3 side-chains (see below).

Building non-canonical non-H3 CDRs

These were built using a modified version (Whitelegg, 1999; N.R.W., Searle & A.R.R., unpublished) of the combined antibody modelling algorithm (CAMAL; Martin *et al.*, 1989). Of special note are the force field modifications in which the stand-alone implementation of VFF (Dauber-Osguthorpe *et al.*, 1988) was used to energy-screen the generated conformations, instead of Eureka. In addition, the candidate conformations were energy-minimised using VFF.

Building H3 loops

Recent studies have shown that in the many instances, CDR-H3 loops exhibit a "kinked" C-terminal region (Shirai *et al.*, 1996; Morea *et al.*, 1998). The kink is determined by a number of sequence-based rules, and the antibodies modelled here conformed to those rules. These features were added to CAMAL and taken into account in the following way when modelling H3 loops. First, the C-alpha to C-alpha database search (Martin *et al.*, 1989) was performed, using only the set of kinked H3 crystal structures to derive the C-alpha to C-alpha distance constraints (in order to bias the database hits in favour of kinked structures). The database hits were then subjected to energy minimisation after grafting, to minimise poor loop/framework joins. Subsequently, the CONGEN (Bruccoleri & Karplus, 1987) and the Go and Scheraga chain closure algorithm (Go & Scheraga, 1970) were used to rebuild a five-residue segment at the apex of each database loop. Finally, each conformation was subjected to minimisation by VFF (Dauber-Osguthorpe *et al.*, 1988) to relieve any residual high energy in the loops, and the lowest energy loops, as measured by VFF, were clustered using RMSD as the criterion.

Final screen

To select the final H3 backbone conformation, a screen based on similarity (measured by RMSD) to the known structures of the particular H3 length being modelled was used. It has been observed that certain length H3 loops have similar conformations in the majority of examples, notably eight and ten residue loops. A similarity score for each of the clustered conformations was calculated: Score $= \Sigma_{\text{all known structs}} (1/d_x)\, m_x$ where d_x is the RMS deviation between the model and structure x, and m_x is the sequence similarity (using the MDM matrix, Schwartz & Dayhoff, 1978) between the sequence

of the model and structure x. The structure being modelled was excluded from the list of known structures while performing this screen. Finally, side-chains were modelled using the dead-end elimination algorithm (see below).

H3 and non-conserved side-chain modelling

The final stage was to build side-chains on the selected backbone model. Analyses of antibodies have shown that for certain residue types, the chi-1 values of the four C-terminal H3 residues are conserved (N.R.W. & A.R.R., unpublished results). Therefore, these were built using the average observed chi-1 values. For the chi-2 of these residues, either the extended conformation was used or the mean of frequently observed (64-88%) values was used, depending on whether the chi-2 was conserved for that residue type and position. All remaining H3 and canonical non-conserved side-chains (as well as all side-chains on the non-canonical non-H3 loops) were built simultaneously on the final selected backbone, using a modified version of the dead-end elimination algorithm (Lasters *et al.*, 1995). In this procedure an attempt is made to find a global minimum by successively eliminating rotamers from a rotamer library which have a higher "lowest" energy of interaction with their environment than the "highest" energy of interaction of the other rotamers for that residue. The Connolly structures (Connolly, 1983) were generated using the program Insight-II 97.0 (MSI) and the images shown are of the solvent-accessible surfaces, as opposed to the van der Waals surfaces.

FACS analysis and immunohistochemistry

Fabs and scFvs were tested by FACS analysis for specific binding to EGP-2 positive cells as described by Henderikx *et al.*. (1998) with some minor modifications. The colon cell line $CaCO_2$ (American Type Culture Collection, Rockville, MD) was stained using scFv (MOC-31) or Fab (clones C3 and 9E) periplasmatic fractions. After the last incubation with FITC-conjugated rabbit anti-mouse (DAKO, Glostrup, Denmark; diluted 1:50 in incubation buffer), detection of bound scFvs or Fabs was performed by means of flow cytometry on a FACScalibur (Becton Dickinson) and results were analysed with the CELLQuest program (Becton Dickinson).

Fabs were tested immunohistochemically as described for scFv by Roovers *et al.* (1998).

Acknowledgements

We thank Dr Hans de Haard for his support and valuable discussions and Professor D. Herlyn for her kind gift of EGP-2. E. van der Linden and M.G.S. Kunen are acknowledged for technical support and expertise. This research was financially supported by the Netherlands Technology Foundation (S.T.W.) and was coordinated by the Earth and Life Sciences Foundation (A.L.W., project MGN55.3858, 805.17.753). N.R.W. and A.R.R. acknowledge support from the Cancer Research Campaign.

References

Adams, G. P., Schier, R., Marshall, K., Wolf, E. J., McCall, A. M., Marks, J. D. & Weiner, L. M. (1998). Increased affinity leads to improved selective tumor delivery of single-chain Fv antibodies. *Cancer Res.* **58**, 485-490.

Baca, M., Presta, L. G., O'Connor, S. J. & Wells, J. A. (1997). Antibody humanization using monovalent phage display. *J. Biol. Chem.* **272**, 10678-10684.

Balaban, E. P., Walker, B. S., Cox, J. V., Sein, A. A., Abrams, P. G., Salk, D., Sheehan, R. G. & Frenkel, E. P. (1991). Radionuclide imaging of bone marrow metastases with a Tc-99m labelled monoclonal antibody to small cell lung carcinoma. *Clin. Nucl. Med.* **16**, 732-736.

Bruccoleri, R. E. & Karplus, M. (1987). Prediction of the folding of short polypeptide segments by uniform conformational sampling. *Biopolymers*, **26**, 137-168.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M. & Poljak, R. J. (1989). Conformations of immunoglobulin hypervariable regions. *Nature*, **342**, 877-883.

Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992). Structural repertoire of the human VH segments. *J. Mol. Biol.* **227**, 799-817.

Connolly, M. L. (1983). Analytical molecular surface calculation. *J. Appl. Crystallog.* **16**, 548.

Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M. & Hagler, A. T. (1988). Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins: Struct. Funct. Genet.* **4**, 31-47.

de Haard, H. J. W., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderikx, P., de Bruine, A. P., Arends, J. W. & Hoogenboom, H. R. (1999). A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J. Biol. Chem.* **274**, 18218-18230.

de Jonge, M. W. A., Kosterink, J. G. W., Bin, Y. Y., Bulte, J. W. M., Kengen, R. A. M., Piers, D. A., The, T. H. & de Leij, L. (1993). Radioimmunodetection of human small cell lung carcinoma xenografts in the nude rat using 111In-labelled monoclonal antibody MOC-31. *Eur. J. Cancer*, **29A**, 1885-1890.

Elias, D. J., Hirschowitz, L., Kline, L. E., Kroener, J. F., Dillman, R. O., Walker, L. E., Robb, J. A. & Timms, R. M. (1990). Phase I clinical comparative study of monoclonal antibody KS1/4 and KS1/4-methotrexate immunoconjugate in patients with non-small cell lung carcinoma. *Cancer Res.* **50**, 4154-4159.

Elias, D. J., Kline, L. E., Robbins, B. A., Johnson, H. C., Jr, Pekny, K., Robb, J. A., Walker, L. E., Kosty, M. & Dillman, R. O. (1994). Monoclonal antibody KS1/4-methotrexate immunoconjugate studies in non-small cell lung carcinoma. *Am. J. Respir. Crit. Care Med.* **150**, 1114-1122.

Figini, M., Marks, J. D., Winter, G. & Griffiths, A. D. (1994). *In vitro* assembly of repertoires of antibody chains on the surface of phage by renaturation. *J. Mol. Biol.* **239**, 68-78.

Figini, M., Obici, L., Mezzanzanica, D., Griffiths, A., Colnaghi, M. I., Winter, G. & Canevari, S. (1998). Panning phage antibody libraries on cells: isolation

of human fab fragments against ovarian carcinoma using guided selection. *Cancer Res.* 58, 991-996.

Foote, J. & Winter, G. (1992). Antibody framework residues affecting the conformation of the hypervariable loops. *J. Mol. Biol.* 224, 487-499.

Go, N. & Scheraga, H. A. (1979). Ring closure and local conformational deformations of chain molecules. *Macromolecules*, 3, 178.

Hawkins, R. E., Russell, S. J. & Winter, G. (1992). Selection of phage antibodies by binding affinity. *J. Mol. Biol.* 226, 889-896.

Helfrich, W., Kroesen, B. J., Roovers, R. C., Westers, L., Molema, G., Hoogenboom, H. R. & de Leij, L. (1998). Construction and characterization of a bispecific diabody for retargeting T cells to human carcinomas. *Int. J. Cancer*, 76, 232-239.

Henderikx, P., Kandilogiannaki, M., Petrarca, C., von Mensdorff-Pouilly, S., Hilgers, J. H. M., Krambovitis, E., Arends, J. W. & Hoogenboom, H. R. (1998). Human single-chain Fv antibodies to MUC1 core peptide selected from phage display libraries recognize unique epitopes and predominantly bind adenocarcinoma. *Cancer Res.* 58, 4324-4332.

Herlyn, M., Steplewski, Z., Herlyn, D. & Koprowski, H. (1979). Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies. *Proc. Natl Acad. Sci. USA*, 76, 1438-1452.

Jespers, L. S., Roberts, A., Mahler, S. M., Winter, G. & Hoogenboom, H. R. (1994). Guiding the selection of human antibodies from phage display repertoires to a single epitope of an antigen. *Biotechnol.* 12, 899-903.

Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. (1986). Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature*, 321, 522-525.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*, 5th edit., U.S. Department of Health and Human Services, Public Health Service National Institutes of Health, Bethesda.

Khazaeli, M. B., Conry, R. M. & LoBuglio, A. F. (1994). Human Immune response to monoclonal antibodies. *J. Immunother*, 15, 42-52.

Kosterink, J. G. W., de Jonge, M. W. A., Smit, E. F., Piers, D. A., Kengen, R. A. M., Postmus, P. E., Shochat, D., Groen, H. J. M., The, H. T. & de Leij, L. (1995). Pharmacokinetics and scintigraphy of Indium-111-DTPA-MOC-31 in small-cell lung carcinoma. *J. Nucl. Med.* 36, 2356-2362.

Kroesen, B. J., Buter, J., Sleijfer, D. T., Janssen, R. A., van der Graaf, W. T., The, T. H., de Leij, L. & Mulder, N. H. (1994). Phase I study of intravenously applied bispecific antibody in renal cell cancer patients receiving subcutaneous interleukin 2. *Brit. J. Cancer*, 70, 652-661.

Lasters, I., de Maeyer, M. & Desmet, J. (1995). Enhanced dead-end elimination in the search for the global minimum energy conformation of a collection of protein side chains. *Protein Eng.* 8, 815-822.

LeMaistre, C. F., Edwards, D. P., Krolick, K. A. & McGuire, W. L. (1987). An immunotoxin cytotoxic for breast cancer cells in vitro. *Cancer Res.* 47, 730-734.

Malby, R. L., Tulip, W. R., Harley, V. R., McKimm-Breschkin, J. L., Laver, W. G., Webster, R. G. & Colman, P. M. (1994). The structure of a complex between the NC10 antibody and influenza virus neuraminidase and comparison with the overlapping binding site of the NC41 antibody. *Structure*, 2, 733-746.

Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. & Winter, G. (1991). By passing immunization: human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.* 222, 581-597.

Martin, A. C., Cheetham, J. C. & Rees, A. R. (1989). Modelling antibody hypervariable loops: a combined algorithm. *Proc. Natl Acad. Sci. USA*, 86, 9268-9272.

Morea, V., Tramontano, A., Rustici, M., Chothia, C. & Lesk, A. M. (1998). Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J. Mol. Biol.* 275, 269-294.

Rader, C., Cheresh, D. A. & Barbas, C. F. (1998). A phage display approach for rapid antibody humanization: Designed combinatorial V gene libraries. *Proc. Natl Acad. Sci. USA*, 95, 8910-8915.

Riethmuller, G., Schneider-Gadicke, E., Schlimok, G., Schmiegel, W., Raab, R., Hoffken, K., Gruber, R., Pichlmaier, H., Hirche, E., Pichlmaier, R., Buggisch, P. & Witte, J. German Cancer Aid 17-IA Study Group (1994). Randomised trial of monoclonal antibody for adjuvant therapy of resected Dukes C' colorectal carcinoma. *Lancet*, 343, 1177-1183.

Riethmuller, G., Holz, E., Schlimok, G., Schmiegel, W., Raab, R., Hoffken, K., Gruber, R., Funke, I., Pichlmaier, H., Hirche, H., Buggisch, P., Witte, J. & Pichlmayr, R. (1998). Monoclonal antibody therapy for resected Dukes' C colorectal cancer: seven-year outcome of a multicenter randomized trial. *J. Clin. Oncol.* 16, 1788-1794.

Roguska, M. A., Pedersen, J. T., Henry, A. H., Searle, S. M., Roja, C. M., Avery, B., Hoffee, M., Cook, S., Lambert, J. M., Blattler, W. A., Rees, A. R. & Guild, B. C. (1996). A comparison of two murine monoclonal antibodies humanized by CDR-grafting and variable domain resurfacing. *Protein Eng.* 9, 895-904.

Roovers, R. C., Henderikx, P., Helfrich, W., van der Linden, E., Reurs, A., de Bruine, A., Arends, J. W., de Leij, L. & Hoogenboom, H. R. (1998). High affinity recombinant phage antibodies directed to the pan-carcinoma marker epithelial glycoprotein-2 for tumor targeting. *Brit. J. Cancer*, 78, 1407-1416.

Salfeld, J. G., Allen, D. J., Kaymakcalan, Z., Labkovsky, B., Mankovich, J. A., McGuiness, B. T., Roberts, A. J., Sakorafas, P., Hoogenboom, H. R. J. M., Schoenhaut, D., Vaughan, T. J., White, M. & Wilton, A. J. (1997). Human antibodies that bind human TNF.agr. patent WO9729131A1.

Schier, R. & Marks, J. D. (1996). Efficient in vitro affinity maturation of phage antibodies using BIAcore guided selections. *Hum. Antibod.* 7, 97-105.

Schier, R., Bye, J., Apell, G., McCall, A., Adams, G. P., Malmqvist, M., Weiner, L. M. & Marks, J. D. (1996). Isolation of high-affinity monomeric human anti-c-erbB-2 single-chain Fv using affinity-driven selection. *J. Mol. Biol.* 255, 28-43.

Schwartz, R. M. & Dayhoff, M. O. (1978). *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed.), vol. 5, pp. 353-362, National Biomedical Research Foundation, Washington DC.

Shirai, H., Kidera, A. & Nakamura, H. (1996). Structural classification of CDR-H3 in antibodies. *FEBS Letters*, 399, 1-8.

Strassburg, C. P., Kasai, Y., Seng, B. A., Miniou, P., Zaloudik, J., Herlyn, D., Koprowski, H. &

Linnenbach, A. J. (1992). Baculovirus recombinant expressing a secreted form of a transmembrane carcinoma-associated antigen. *Cancer Res.* **15**, 815-821.

Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992). The repertoire of human germline VH sequences reveals about fifty groups of VH segments with different hypervariable loops. *J. Mol. Biol.* **227**, 776-798.

Tomlinson, I. M., Cox, J. P. L., Gherardi, E., Lesk, A. M. & Chothia, C. (1995). The structural repertoire of the human V kappa domain. *EMBO J.* **14**, 4628-4638.

Watzka, H., Pfizenmaier, K. & Moosmayer, D. (1998). Guided selection of antibody fragments specific for human interferon gamma receptor 1 from a human VH- and VL-gene repertoire. *Immunotechnology*, **3**, 279-291.

Whitelegg, N. R. (1999). Molecular modelling of antibody combining sites, *thesis*.

Willems van Dijk, K., Mortari, F., Kirkham, P. M., Schroeder, H. W. & Milner, E. C. B. (1993). The human immunoglobulin VH7 gene family consists of a small, polymorphic group of six to eight gene segments dispersed throughout the VH locus. *Eur. J. Immunol.* **23**, 832-839.

Williams, S. C., Frippiat, J. P., Tomlinson, I. M., Ignatovich, O., Lefranc, M. P. & Winter, G. (1996). Sequence and evolution of the human germline V lambda repertoire. *J. Mol. Biol.* **264**, 220-232.

Zimmermann, S., Wels, W., Froesch, B. A., Gerstmayer, B., Stahel, R. A. & Zangemeister-Wittke, U. (1997). A novel immunotoxin recognising the epithelial glycoprotein-2 has potent antitumoural activity on chemotherapy-resistant lung cancer. *Cancer Immunol. Immunother.* **44**, 1-9.

Edited by I. A. Wilson

(Received 7 October 1999; received in revised form 6 January 2000; accepted 6 January 2000)